

深度学习方法在糖尿病视网膜病变诊断中的应用

范家伟¹ 张如如^{1,2} 陆萌^{1,2} 何佳雯^{1,2} 康霄阳^{1,2} 柴文俊^{1,2}
石坤达^{1,2} 宋美娜^{1,2} 鄂海红^{1,2} 欧中洪^{1,2}

摘要 深度学习可以有效提取图像隐含特征,在医学影像识别方面的应用快速发展. 由于糖尿病视网膜病变 (Diabetic retinopathy, DR) 诊断标准明确、分类体系成熟,应用深度学习诊断糖尿病视网膜病变近年来成为研究热点. 本文从深度学习在 DR 诊断中的最新研究进展、DR 诊断的一般流程、公共数据集、医学影像标注方法、主要实现模型、面临的主要挑战几方面,对深度学习方法在糖尿病视网膜病变诊断中的应用进行了详细综述,便于更多机器视觉、尤其是深度学习医学影像的研究者们参照对比,加快该领域研究的成熟度和临床落地应用.

关键词 深度学习, 糖尿病, 糖尿病视网膜病变, 智能诊断, 图像标注, 病变区域检测, 病变等级分类

引用格式 范家伟, 张如如, 陆萌, 何佳雯, 康霄阳, 柴文俊, 石坤达, 宋美娜, 鄂海红, 欧中洪. 深度学习方法在糖尿病视网膜病变诊断中的应用. 自动化学报, 2021, 47(5): 985-1004

DOI 10.16383/j.aas.c190069

Applications of Deep Learning Techniques for Diabetic Retinal Diagnosis

FAN Jia-Wei¹ ZHANG Ru-Ru^{1,2} LU Meng^{1,2} HE Jia-Wen^{1,2} KANG Xiao-Yang^{1,2} CHAI Wen-Jun^{1,2}
SHI Shen-Da^{1,2} SONG Mei-Na^{1,2} E Hai-Hong^{1,2} OU Zhong-Hong^{1,2}

Abstract Deep learning can effectively extract the hidden features of image and its application in medical image recognition is developing rapidly. Due to the clear diagnostic criteria for diabetic retinopathy (DR) and the mature classification system, the application of deep learning to diagnose diabetic retinopathy has become a research hot-spot in recent years. Therefore, this paper reviews the application of deep learning methods in the diagnosis of diabetic retinopathy detailedly based on the latest research progress of deep learning for DR diagnosis, the general flow for DR diagnosis, public dataset, medical image annotation method, main models and major challenges. It brings convenience for more researchers of computer vision deep learning, especially medical imaging deep learning, to speed up the research maturity and clinical application in this field.

Key words Deep learning, diabetes mellitus (DM), diabetic retinopathy (DR), intelligent diagnosis, image annotation, lesions detection, lesions classification

Citation Fan Jia-Wei, Zhang Ru-Ru, Lu Meng, He Jia-Wen, Kang Xiao-Yang, Chai Wen-Jun, Shi Shen-Da, Song Mei-Na, E Hai-Hong, Ou Zhong-Hong. Applications of deep learning techniques for diabetic retinal diagnosis. *Acta Automatica Sinica*, 2021, 47(5): 985-1004

近年来,深度学习在各种计算机视觉任务中的突出表现,推动了其在医学影像识别中的发展,使深度学习技术处理医学影像成为一个重要的研究方向.深度学习技术利用医学影像数据在多个领域如乳腺病变诊断、皮肤病变诊断、肺癌诊断以及糖尿病视网膜病变 (Diabetic retinopathy, DR) 诊断等

方面已取得了医生级别的准确率.

DR 是糖尿病 (Diabetes mellitus, DM) 的重要并发症之一,据统计,全球约有 4.25 亿人患有糖尿病,目前,患病人数仍在激增,预计到 2035 年,将会达到 5.92 亿^[1]. 调查显示,其中约 1/3 的糖尿病患者将发生 DR,近 10% 的 DM 患者将发生威胁视力的视网膜病变 (Vision-threatening diabetic retinopathy, VTDR),糖尿病视网膜病已成为全球工作人群失明人数攀升的主要原因. 由于 DR 检查量增加,在诊断方面出现了以下问题: 1) 大量的阅片任务,给医生带来了极大的工作压力; 2) 医生阅片速度慢,患者也无法得到即时的反馈; 3) 优质医疗资源被占用,漏诊、误诊人数增长^[2]. 因此,DR 早期诊断面临严峻的挑战.

收稿日期 2019-01-28 录用日期 2019-06-09
Manuscript received January 28, 2019; accepted June 9, 2019
教育部信息网络工程研究中心资助项目资助
Supported in part by Engineering Research Center of Information Networks of Ministry of Education
本文责任编辑 张道强
Recommended by Associate Editor ZHANG Dao-Qiang
1. 北京邮电大学计算机学院 北京 100089 2. 教育部信息网络工程研究中心 (北京邮电大学) 北京 100089
1. School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100089 2. Information Network Engineering Research Center, Ministry of Education (Beijing University of Posts and Telecommunications), Beijing 100089

国际 DR 诊断标准的全球公认性以及分类明确性为 DR 智能诊断系统的研发提供了统一标准与基础, 所以更多的研究者以及公司企业将目光投向这一领域. 经过不断优化, DR 智能诊断系统逐渐达到了专家水平, 不仅可以判断患者是否存在 DR, 还可以对患者的病变等级进行详细划分, 并对眼部的不同病理特征区域进行检测标识, 提高了系统的可解释性. 部分 DR 诊断系统的输出结果不再只是给医生做决策支持, 而是可直接为患者做出精准的诊断, 提供更详细的信息及诊断依据^[3]. 另外, 目前已经有基于深度学习的 DR 诊断系统上市用于临床实践, 这代表该项技术有很强的实用性与落地性, 为 DR 智能诊断系统的研究带来革命性的进展.

本文的主要贡献如下: 对深度学习方法在糖尿病视网膜病变诊断中的应用进行了全面梳理; 对各类公共数据集、医学影像标注方法、病变区域分割及检测模型、病变等级分类模型以及模型评估方法进行了综述; 对 DR 病变区域检测、病变等级分类的最新进展进行了归纳; 对比了各种实现方法的优缺点. 通过阅读本文, 可以获悉该领域的发展现状、未来发展方向以及面临的挑战, 便于研究者参照对比, 加快该领域的研究及临床落地应用.

1 相关知识

1.1 糖尿病视网膜病变病理及诊断标准

糖尿病视网膜病变由糖尿病微血管病变导致, 大致可分为血管破裂出血、释放生长因子、血管堵塞三个环节. 当人体血糖过高, 会引发微血管基底膜增厚, 进而血管口径减小、内壁变粗糙、弹性和收缩力减弱; 此时, 分布在视网膜上的微血管由于十分脆弱, 将极易破裂、出血, 并释放血管内皮生长因子 (Vascular endothelial growth factor, VEGF);

血管释放出的 VEGF 会刺激临近的新形成的毛细血管, 后者由于管壁极薄, 受到轻微刺激便会破裂, 上述这几个环节循环往复, 使得视网膜受损愈加严重^[4].

糖尿病视网膜病变患者在患病的不同阶段会出现不同的病理特征, 主要包括: 微动脉瘤、出血和渗出物 (包含硬渗出和软渗出). 视网膜微动脉瘤出现在糖尿病视网膜早期, 同时伴随有视网膜血管异常漏血. 渗出包括硬性渗出和软性渗出两种, 硬性渗出多由扩张的毛细血管和微血管瘤渗漏的脂质和蛋白质成分组成, 沉积于外层视网膜和视网膜下形成边缘明确的斑; 软性深渗出则临床表现为形状不规则、边界模糊、大小不等的棉絮或绒毛样网膜渗出斑^[5]. 若保持长期的病理状态, 部分血管会出现堵塞或闭合, 以至营养物质无法送至视网膜, 造成眼部的大面积损伤, 最终导致失明.

根据上述病理特征, 糖尿病视网膜病变分为两个阶段^[6]:

1) 非增殖性糖尿病视网膜病变 (Non proliferative diabetic retinopathy, NPDR). NPDR 是 DR 的早期阶段. NPDR 的诊断有助于对疾病进展和视力丧失进行风险预测, 并确定随访的时间间隔.

2) 增殖性糖尿病视网膜病变 (proliferative diabetic retinopathy, PDR). PDR 是 DR 的严重阶段, 体现了广泛视网膜缺血和毛细血管闭锁导致的血管生长反应.

在糖尿病视网膜病变的研究中, 为了更好地地区分患者糖尿病视网膜病变的严重程度, 便于更明确地界定病情进展^[7], 通常采用分期标准. 现行最通用的是国际五级诊断标准. 该分类方法按照病症进展情况分为五期: I 期 (无病症)、II 期 (轻度非增殖期)、III 期 (中度非增殖期)、IV 期 (重度非增殖期)、V 期 (增生期), 相应诊断标准如表 1 所示^[8].

表 1 糖尿病视网膜病变国际分级标准

Table 1 International classification of diabetic retinopathy and diabetic macular edema

| 糖尿病视网膜病变 | 散瞳后眼镜所见 | 随访建议 |
|-------------------------|---|-------------------|
| 无视网膜病变 (无DR) | 无异常 | 1 ~ 2 年随访一次 |
| 轻度非增殖性糖尿病视网膜病变 (轻度NPDR) | 仅有微动脉瘤 | 1 ~ 2 年随访一次 |
| 中度非增殖性糖尿病视网膜病变 (中度NPDR) | 比仅有微动脉瘤重, 比重者轻 | 半年到一年随访一次或转诊至眼科医师 |
| 重度非增殖性糖尿病视网膜病变 (重度NPDR) | 有以下任一症状之一: 4个象限每个都有20个以上的内出血病灶; 2个以上象限有确定的静脉珠状改变; 1个以上象限有明显的视网膜内微血管异常并且无增殖性病体征 | 转诊至眼科医师 |
| 增殖性糖尿病视网膜病变 (PDR) | 具有重度非增殖性症状且有以下一种或多种情形: 新生血管, 玻璃体积血/视网膜前出血 | 转诊至眼科医师 |

1.2 成像设备与技术

不同的成像设备拍摄出的影像数据形式不同, 从而会使影像标注方法与深度学习算法的性能产生差异. 现阶段应用范围最广的两种成像技术为彩色眼底成像技术 (Color fundus photography, CFP) 以及光学相干断层扫描 (Optical coherence tomography, OCT) 技术.

1.2.1 眼底成像技术

通常眼底成像技术通过眼底照相机来实现, 近年来, 眼底成像技术的发展, 为 DR 诊断提供了高分辨率的影像数据: 数据中病理特征越来越明确, 病变区域轮廓越来越清晰. 这不仅为标注工作提供了方便, 也大大提高了 DR 智能诊断系统的性能.

1) 无赤光技术. 在眼底成像中, 让光线通过特制的滤光片, 以达到过滤短波光线的效果. 应用该技术拍摄出的眼底影像会增加视网膜血管和背景的对比度, 也可以很容易地通过颜色区分脉络膜与视网膜损害^[9]. 由于该类影像所呈现的病理特征明确, 不同病理特征区分度明显, 使用该类影像作为深度学习数据集可以有更好的效果.

2) 立体眼底成像技术. 眼底相机通过分光镜或者通过不同角度的两次拍摄, 生成左右并列的图像再通过立体镜识图, 得到立体成像^[10]. 该类影像包含隆起与凹陷的特征信息, 在进行深度学习时, 可以通过该类信息进行更好的病征识别.

3) 共聚焦激光扫描技术. 全景数码摄像设备通过共聚焦光学原理, 拍摄出更大视角的眼底图像. 这有利于检测位于边缘的病变特征.

另外, 眼底镜的镜头、光源特性等都会影响成像的效果, 以至于影响深度学习的效果.

1.2.2 光学相干断层扫描技术

光学相干断层扫描技术是一种新型层析成像技术, 能无损、快速地获得样品的高分辨断层图像, 它可以对材料及生物系统内部微观结构进行高分辨率横断面层析成像^[11]. 它在眼科领域与传统成像技术相比有以下几个优势:

1) 能获得人眼内各个断层的图像, 增大了医师所获取的信息量;

2) 具有更高的分辨率, 医师可以对眼部某一特定部位进行研究和分析^[12];

3) 传统的眼底相机成像利用的是光的阻断特性, 而 OCT 是利用了光的散射特性, 这可以获取与传统眼底相机不同的信息;

4) 成像设备体积小;

5) 短时间内可获取大量数据^[13].

通过深度学习技术对 OCT 影像进行分析, 可以更精确地捕捉到患者病情信息, 做出更准确的诊断.

1.2.3 成像设备的发展趋势

从整体上来说, 成像设备的发展有两大趋势: 高技术化与便携化. 高技术的成像设备可提供更加清晰、有效的影像, 辅以医生诊断, 例如 OCT 的出现给患者提供了更加优质的检查. 便携化的成像设备可以为医生和患者提供更加方便的服务, 随着在线医疗远程医疗的发展, 便携化设备的作用会进一步凸显: 患者通过手机等智能设备完成眼底检查^[14], 并上传至平台, 医生通过软件对患者病情进行实时追踪, 这不仅会省去繁琐的检查步骤、节省大量的时间成本, 而且患者可以实时得到反馈, 提高了效率. 例如印度 Remidio 公司开发了一款基于手机的眼底相机, 通过软件医生可看到患者的眼底情况, 以观察患者病情进展. 这是一次在该领域的有意义的尝试.

2 基于深度学习的 DR 智能诊断系统发展现状

2.1 最新发展现状

基于深度学习技术对 DR 进行早发现、早干预、早治疗可以有效降低患者致盲风险, 引起了学术界和产业界的广泛关注.

2.1.1 学术界进展

2018 年 2 月, 美国德克萨斯大学研究者^[15]将迁移学习算法用于 OCT 影像诊断, 在 ImageNet 数据集上训练 InceptionV3 网络, 进行微调后用于 OCT 影像训练, 达到较高诊断精度, 解决了小数据集的问题. 同年 10 月广州中山大学中山眼科中心眼科学国家重点实验室联合多家研究机构共同研发了基于中国人群彩色眼底影像 DR 检测的深度学习算法 (Deep learning algorithm, DLA), 并且该算法也在其他种族人群中进行验证, 取得了较高的精确度^[16-17]. 2018 年 8 月, 谷歌公司 DeepMind 团队提出了一种两阶段深度神经网络模型用于 OCT 影像诊断, 分别实现了病灶区域分割和病变等级分类, 并通过不同厂商的 OCT 成像设备获取的影像进行验证, 准确率高达 95%^[18].

2.1.2 产业界进展

2017 年 9 月杭州市卫计委在全国率先启动糖尿病视网膜病变筛查项目, 引入了基于人工智能辅助诊断的手持式眼底照相机, 解决了基层眼科专业人员匮乏的问题. 2018 年 4 月, 世界首个获批用于

DR 诊断的人工智能 (Artificial intelligence, AI) 产品 IDx-DR 美国上市^[19], 可用于检测 22 岁及以上 DM 患者的视网膜病变程度 (包括黄斑水肿), 并且不需要专门的医生对检测报告进行解读. 2018 年, 印度 Remidio 公司研发出了高质量便携式视网膜成像设备 FOP (Fundus on phone), 只需配合智能手机以及人工智能算法 EyeArt 就可以进行详细的眼部健康筛查^[20]. 2018 年 6 月, 中国第一个投入使用的 AI 医疗机器人——“嵩岳”医生^[21], 进驻河南省人民医院, 提高了 DR 诊断和筛查水平, 可及时发现“病灶”解决问题.

2.2 基于深度学习 DR 诊断的一般框架

DR 诊断的一般框架包括数据获取、医学影像标注、病灶区域检测、病变等级分类及模型评估等步骤. 如图 1 所示.

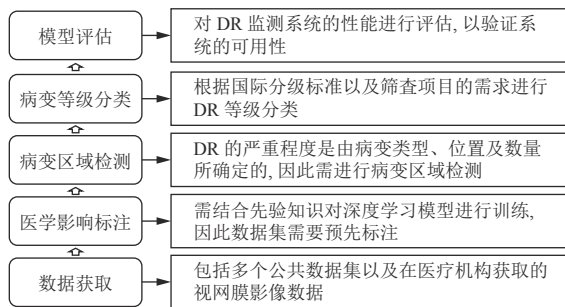


图 1 基于深度学习 DR 诊断的一般框架

Fig.1 General framework of diabetic retinal diagnosis based on deep learning

DR 诊断框架流程介绍:

1) 数据获取. 数据是深度学习算法所需的核心资源, 为了训练和测试不同 DR 检测算法, 需要大量的眼底影像, 主要包括公共数据集以及医疗机构获取数据. 在 DR 检测领域已存在多个公共数据集, 便于研究者直接用来训练测试模型, 促进 DR 智能诊断系统的研究与发展.

2) 医学影像标注. 深度学习需要结合先验知识对模型进行训练, 从医疗机构获取的影像数据需要预先标注, 而多数情况下依赖专家手工标注, 耗费大量人力物力, 因此如何有效标注海量医学影像数据成为智能诊断系统的挑战之一, 需要更先进更可靠的医学影像标注方法进行改进和优化.

3) 病变区域检测. DR 的严重程度是由病灶 (如微动脉瘤、出血、渗出物等) 的类型、位置及数量所确定的, 但提取这些病灶特征对传统医生或临床医生来说是一项耗时的任务^[22], 因此需要搭建强大的病灶区域检测系统, 以提取复杂的病灶特征.

4) 病变等级分类. 在世界各地, 不同 DR 筛查

项目使用不同的 DR 分类系统^[23], 根据筛查项目的需求、视网膜影像中病灶信息以及国际 DR 诊断标准进行不同病变等级分类, 主要包括是否患有 DR、是否需要转诊、是否患有威胁视力的 DR 以及五等级划分等分类方法.

5) 模型评估. 对 DR 检测系统的性能进行评估, 以验证系统的可用性, 本文中介绍了几种常用的评估 DR 检测算法的性能指标, 包括准确性、敏感度 (召回率)、特异度、精确度、F-score、ROC (Receiver operating characteristic) 曲线等.

3 数据集及医学图像标注

3.1 公共数据集

获得训练样本最直接的方式是直接使用公共数据集, 不同数据集在仪器设备、拍摄角度、患病情况、标注情况、图像样式、图像尺寸等方面都有较大的区别, 开发者可根据自己不同的需求选用合适的数据集. 详细内容如表 2 所示.

3.2 数据集的使用及数据预处理

在开发过程中, 通常会使用多种数据集. 因此解决不同源数据集的成像差异成为一个重要的问题. 差异主要表现为两点, 即设备差异和人种差异.

1) 设备差异

由于不同数据源所使用眼底设备不同, 在成像分辨率、成像广度、色彩还原能力等方面有所差异. 因此为消除此类差异, 需要对图像进行预处理, 大体分为以下两步: a) 尺寸归一化; b) 图像增强. 尺寸归一化可以将图像处理成相同规格; 图像增强技术能够增强病理特征, 减弱不同成像设备光照强度、色彩等方面的差异. 下面重点对应用在 DR 诊断上的图像增强做简要介绍.

现阶段图像增强技术已经发展的较为完善, 例如以灰度变换、直方图均衡, 空域滤波为代表的空域增强技术; 以高斯滤波、巴特沃思滤波为代表的频域增强技术.

灰度变换通过对图像三个通道的灰度值分别进行线性或非线性变换, 增大相邻像素的对比差. 直方图均衡通过对灰度直方图的变换, 达到增大局部对比度的效果, 这些方法都达到了增强病症区域与正常区域区别的效果, 使得模型可以获取更加明确的特征. 频域增强技术通过低通滤波器, 过滤掉高频的噪声, 在图像中表现为去掉了高亮的光斑, 使得图像更加平滑.

2) 人种差异

不同人种的视网膜颜色特征、结构特征几乎无

表 2 糖尿病视网膜病变公共数据集
Table 2 Public data set on diabetic retinal

| 名称 | 数据类型 | 图像信息 | 图片尺寸 | 图片格式 | 数据量 (幅) | 获取权限 | 链接 |
|--|------|--|--|------|---------|--------------|---|
| IDRID ^[24] | CFP | 图像按照国际标准进行糖网和黄斑水肿分级, 并对其中 81 幅有糖网征象的图像进行了病变的像素级标注. | 4228 × 2848 | JPEG | 516 | 注册 IEEE 账号获取 | https://iee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid |
| E-Ophtha ^[25] | CFP | 由 e-ophtha-MA (微动脉瘤) 和 E-Ophtha-EX (渗出) 两个子数据库组成. 标注了 EX 和 MA 的区域, 以掩模的方式给出. | 2544 × 1696 1440 × 960 1504 × 1000 等 | JPEG | 463 | 提交邮箱以获取下载码 | https://iee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid |
| DRiDB ^[26] | CFP | 每幅图像由至少 5 名专家标注, 标注内容包含所有视网膜主要解剖结构和病理特征以及糖网分级. | 720 × 676 | BMP | 50 | 发送邮件来请求访问数据库 | https://ipg.fer.hr/ipg/resources/image_database |
| Messidor ^[27] | CFP | 图像标注了糖尿病性视网膜病变分级及黄斑水肿分级. | 1400 × 960 2240 × 1488 2304 × 1536 | TIFF | 1200 | 提交邮箱以获取下载码 | http://www.adcis.net/en/Download-Third-Party/Messidor.html |
| Messidor-2 ^[28] | CFP | 每幅图像标注了糖尿病性视网膜病变分级及黄斑水肿分级. | 1400 × 960 2240 × 1488 2304 × 1536 | TIFF | 1784 | 提交邮箱以获取下载码 | http://latim.univ-brest.fr/indexfce0.html |
| DIARETDB0 ^[29] | CFP | 该数据集是用于糖尿病视网膜病变检测的基准的公开数据集, 每幅图像标注了病变信息. | 1500 × 1152 | PNG | 130 | 直接下载 | http://www.it.lut.fi/project/imageret/diaretdb0/index.html |
| DIARETDB1 ^[30] | CFP | 该数据集可用做基准糖尿病视网膜病变检测数据集, 由 4 名医学专家对糖网相关病变区域进行标注. | 1500 × 1152 | PNG | 89 | 直接下载 | http://www.it.lut.fi/project/imageret/diaretdb1/ |
| Large dataset of OCT on Mendeley ^[15] | OCT | 图像分为训练集和测试集, 每个集合含有 4 种标签的数据: CNV, DME, DR-USEN 和 NORMAL. | 不统一约为 400 × 500 | JPEG | > 50000 | 直接下载 | https://data.mendeley.com/archiver/rschjbr9sj?version=3 |
| Dataset for OCT classification ^[31] | OCT | 由 50 个正常、48 个干性 AMD 和 50 个 DME 组成. | 765 × 765 | TIFF | 3700 | 直接下载 | https://sites.google.com/site/hosseinrabbani/horasgani/datasets-1 |

解剖学差异, 其差异主要体现为虹膜的颜色. 如果期望在不同人种都有更好的泛化能力, 可以通过增加各个人种的眼底数据集来实现. 中山大学团队^[16]搜集了不同国家和地区的不同人种的眼底照片, 标注后用于模型的训练, 该模型在各个人种上都表现出了很好的泛化能力.

3.3 图像标注

图像标注是医学影像智能诊断系统落地过程中至关重要的一环, 它决定了 AI 算法的性能上限. 由于公共数据集数据量相对不足, 在训练过程中难以满足开发者的需求. 因此如何获得高质量的数据集, 如何科学合理的安排和进行标注成为首要问题^[32], 以下将介绍常用的标注工具以及标注流程.

3.3.1 标注工具

现行的标注结果可大致分为三种, 分别为: 基础类别标注; 类别标注+病灶方框标注; 类别标注+像素级勾画. 随着标注的精细度逐级提升, 标注工具也愈加强大^[33]. 下面介绍几种常用的标注方式.

1) PC 机 + 电子病历. 医生在 PC 端根据电子

病历记录, 将存在不同病理特征的影像放到不同的文件夹中, 此类方法工作量较小, 适用于分类模型的训练, 不适用于病灶区域的分割.

2) 专业标注工具. 医师利用标注工具, 如 LabelMe, Labelimg, Pixorize 和 Labelbox 或一些成熟的网站系统, 如基于 H5 及缩放的 Web 标注, 基于 Surface 的画笔标注以及基于登录模式的多人协助标注平台进行标注等, 但由于病灶区域较多以及轮廓不规则等问题, 即使是熟练的医师标注也要耗费大量时间.

3) AI 辅助 + 专业标注工具. 将一些成熟的 DR 诊断系统 (如 EyeWisdom、IDx-DR、EyeArt) 或自主研发的 DR 病灶区域检测和病变等级分类算法集成到专业标注工具中, 搭建基于人工智能的医学影像标注系统. 在标注前, 首先通过标注系统对视网膜影像进行病灶区域以及病变等级预标注, 再由专家通过微调获得标注结果. 此种方法可大大提高医生的标注效率, 也是未来医学影像标注的发展趋势.

3.3.2 标注流程

由于每位专家可能只对一种或几种病灶特征比较熟悉,并且专家标准略有差异,因此标注过程需要统一的标准化流程以及经验丰富的专家团队,以下列举了三种标注流程,供读者参考:

1) 分级标注. 标注人员先利用相应的标注工具进行标注,再由专家审核和补充. 将标注人员分为若干层级,依次执行筛选、标注、审核. 可由经过专家培训和指导的青年医师负责影像的筛选和标注,要明确标注方式、标注病灶类别、标注范围等. 然后由病理学专家负责定期和最终审核,适时修正遗漏或错误的标注;如果依然不能确定,可咨询其他专家予以指导. 2018年 Kermany 等就是采用此类方法进行分级标注^[15],第1级评分者为受过培训的医学本科生,主要负责质量筛选,排除有严重伪影或分辨率低的影像;第2级评分员由四名眼科医生组成,他们对每幅图像进行独立标注;第3级由两位超过20年临床经验的专家组成,验证每幅视网膜影像标签的准确性.

2) 交叉标注. 将视网膜影像随机分给每位医生,通过标注工具对给定影像进行标注分级后返回,再随机分给其他医生,直到得到 N 个(自设)一致的标注结果,并将该结果作为最终影像标签. 在此过程中,标注者无权查看其他人的标注结果,且任何影像只可获得一次. 2018年中山大学研究团队就是应用了此类方法,当实现三名医生标注结果一致时确定为最终标注结果^[16-17].

3) 众包标注. 众包是指将以往在传统方式下由企业或组织内部员工执行的工作任务,以自由自愿的形式借助于网络外包给非特定的大众的做法. 搭建众包标注系统,通过招募及简单的培训测试,筛

选出合格的标注人员之后,将待标注影像分发给标注人员,在质量控制方法的调控下,保证众包标注数据的质量,并接收标注结果. 众包标注有许多质量控制方法,如:多数表决(Majority vote, MV)、DS算法、GLAD算法、贝叶斯算法(RY)、Zen-Crowd算法、AVNC算法等^[34].

4 糖尿病视网膜病灶区域检测

DR 病灶区域检测对 DR 等级分类以及提高诊断系统的可解释性尤为重要. 本节基于各种 DR 病灶特征(渗出物、微动脉瘤、出血等)的检测任务,介绍了几种常用的图像语义分割模型,随之对迁移学习的应用方法进行简单概述,缓解了训练数据量少的问题;最后对 DR 病灶区域检测的相关研究进行综述,了解最新研究进展,并对其优缺点进行总结,具体如图2所示.

4.1 病灶区域分割模型

近年来,越来越多的分割方法用于检测 DR 病灶区域,并取得了较好的效果. 其主要分为两种:第1种是 Encoder-decoder 架构模型,该类架构中最具代表性的是全卷积网络(Fully convolutional network, FCN)^[35]、SegNet^[36]、U-Net^[37]等;第2种是基于 Atrous convolution 的模型,该类架构中最具代表性的是 DeepLab 系列模型,也是目前主流模型.

4.1.1 Encoder-decoder 架构

对糖尿病视网膜病灶区域检测所应用的模型大多是以 Encoder-decoder 架构为基础. 编码器通常是以分类网络如 VGGNet^[38]、ResNet^[39]等进行下采样得到特征图(Feature map);解码器对特征图进行上采样恢复目标细节和相应的空间维度,最后将

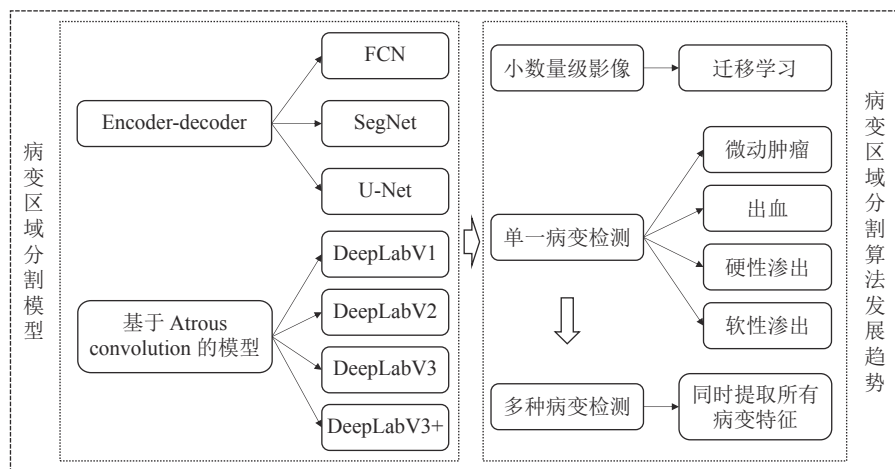


图2 糖尿病视网膜病灶区域检测

Fig. 2 Regional detection of diabetic retinopathy

结果与标注影像像素进行匹配, 反复训练调整参数. 此处介绍两种比较流行的网络架构.

1) U-Net 模型

U-Net^[37] 网络的提出起源于生物医学领域的分割, 并在 2015 年 ISBI (International symposium on biomedical imaging) 神经元结构分割挑战赛中取得了更加优异的性能.

U-net 网络主要由 contracting path 和 expanding path 两部分组成, contracting path 主要用来捕捉影像中上下文信息, 与之对应的 expanding path 则是对影像中分割区域进行精准定位, 将下采样过程中得到的特征图, 拼接至对应上采样阶段特征图上, 并进行信息融合, 从而形成一个梯形结构. 通过这种跳远拼接架构, 在每个阶段都允许解码器学习在编码器池化中丢失的相关特性, 实现对医学影像的精确分割. 目前, 该网络在视网膜病灶区域分割研究方面也得到了广泛的应用, 例如, 2018 年 8 月, 谷歌 DeepMind 团队^[18] 通过 3D U-Net 网络^[37] 实现了对视网膜 OCT 影像中 15 种病变区域的分割, 取得了国际领先水平.

2) SegNet 模型

2017 年 Badrinarayanan 等^[36] 提出了 SegNet 分割模型, 该模型在节省内存和计算量方面都取得了较大的提高. SegNet 模型最大特点在于它的池化层多了一个索引 (Index) 功能, 在进行最大池化时, 选择并记录最大像素在特征图像中的位置, 传递到对应的上采样层, 依据 Index 复原最大像素到原来位置, 其他位置补零, 并与上采样后的特征图进行信息融合. 这样做的好处是: a) 提高边界划分精度; b) 上采样阶段不需要再进行学习, 减少训练参数、节省内存; c) 可以广泛地应用在其他 Encoder-decoder 结构中.

SegNet 与 U-Net 结构相似, 但是有两点显著区别: a) SegNet 模型是将最大池化层拼接至上采样后的特征映射图, 而 U-Net 网络将整个特征映射图送到解码器, 因此 SegNet 会更加节省资源; b) SegNet 可以利用 VGG16 所有预训练卷积层的权重, 而 U-Net 由于内存问题无法使用更深的模块.

因此, 很多研究者将 SegNet 模型应用于糖尿病视网膜病灶区域分割, 如 2018 年 Badar 等^[40] 在 MIUA (Medical image understanding and analysis) 国际比赛中, 通过 SegNet 模型对视网膜渗出物、出血等病灶进行分割, 取得了较高的精确度.

4.1.2 基于 Atrous convolution 的模型

基于 Atrous convolution 的模型解决了训练过程中特征分辨率降低、图像多尺度、卷积模型平移不变性的问题, 在糖尿病视网膜病灶区域检测中,

更是需要分辨率高、特征稠密、定位精确的特征图像来检测 DR 早期症状. DeepLab 模型就是应用了 Atrous convolution^[41] 的思想, 并且在 2018 年 4 月 ISBI 糖尿病视网膜病变—分割和分级挑战赛中, 前三名的团队均采用 DeepLab 模型. 下面对 DeepLab 模型的发展进行简单介绍.

DeepLabV1^[42] 基于 VGG16 模型结合了 Atrous 算法和条件随机场 (Conditional random fields, CRFs), 通过 Atrous 算法扩展感受野缓解下采样或池化降低分辨率的问题, 获取更多的上下文信息; 采用完全连接的 CRF 提高模型捕获细节的能力, 缓解因平移不变性造成的定位精度低的问题.

DeepLabV2^[43] 是对 DeepLabV1 的优化, 该模型主要基于 ResNet 网络, 在最后几个最大池化层中使用 Atrous convolution 取代下采样, 获得高密度的特征映射图. DeepLabV2 提出了空洞空间金字塔池化 (Atrous spatial pyramid pooling, ASPP), 在给定图像上以不同采样率的 Atrous convolution 并行采样, 从而解决多尺度问题, 并且不会增加运算量, 虽然采用 U-Net 和特征金字塔网络 (Feature pyramid network, FPN)^[44] 的思想, 也整体提高了系统的性能, 但同时也增加了特征计算量和存储空间.

DeepLab 延续到 DeepLabV3^[45] 探讨了不同结构方向. DeepLabV3 采用多采样率 Atrous convolution 级联或并行来捕获多尺度背景. 此外, DeepLabV3 改进了 ASPP 模块, 将全局背景基于图像层次进行编码获得特征.

DeepLabV3+^[46] 模型集合了 DeepLabV1-V3 的优点, 为了融合多尺度信息, 引入语义分割常用的自动编码模型, 并在自动编码模型中引入 Xception^[47], 其思想是引入更多的路径, 甚至让每个通道都具有 1 条独立路径, 卷积核独立运作在各自的通道中, 减少参数, 提高运算速度. 通过引入 ASPP 和 Xception 模型思想, 提高 Encoder-decoder 网络的运行速率和健壮性.

4.2 迁移学习

训练神经网络需要大量的标注数据, 而医学领域数据集的构建成本非常高, 因此需要一种在有限数据资源下满足神经网络训练的方法.

其中, 解决这个问题的一种方法是数据增强, 通过对图像进行旋转、平移、裁剪等达到增大数据集的目的, 但是迄今为止使用的数据增强技术并没有创建真正的样本, 所以有待继续发展. 另外一种常用的有效的方法就是直接迁移学习^[48]. 直接迁移学习是一种利用海量数据集对神经网络进行训

练,然后将训练后的网络进行微调,应用于其他分割和分类任务.由于在医学诊断中,大多病灶区域有相对明显的边缘特征,因此可以通过基于特征迁移学习达到良好的效果.

除了直接迁移学习外,常用的利用迁移学习技术解决医学领域问题的方法还包括半监督迁移和跨域适应(Domain adaptation)等方式.半监督迁移学习旨在解决源域和目标域数据不平衡的问题,其主要思想是使用源域的丰富数据和标签帮助解决目标域因数据稀少造成的性能差的问题^[49].跨域适应旨在从信息丰富的源数据分布中学习,提升在相关但不同的目标数据分布上的模型性能,Dou等^[50]提出了一种具有对抗学习的无监督的跨域适应框架,用于跨模态的生物医学图像分割.

如今,迁移学习已经广泛应用于眼底疾病的诊断中,在视网膜病变检测的研究上也有很多突破性进展.最具代表性的就是Keremany等^[15]、Shan等^[51]将迁移学习应用于视网膜OCT检测,并取得了较好的效果.

4.3 病灶区域检测相关研究

针对不同类型的DR病灶渗出物(包括:硬性

渗出物(Hard exudate, HE)和软性渗出物(Soft exudate, SE)、微动脉瘤(Microaneurysm, MA)、出血(Hematoma expansion, HE)等,研究者提出了不同的基于深度学习进行DR病灶区域检测和分割的方法,在本小节中,我们将对这些方法进行回顾,如表3所示.

通过对以上文献的回顾表明,随着深度学习技术的发展,越来越多的病灶区域检测方法涌现出来.其中,Shan等^[51]通过生产MA patches的方法检测MA,取得了较好的效果,但Patch大小的选择对模型性能具有一定的影响,需要对模型进行反复调参;Abbasi-Sureshjani等^[52]与Shan等^[51]一样采用Patches检测方式,但着重选择具有高度误导性的非渗出物样本作为Patches,达到网络自动学习调整参数的目的;van Grinsven等^[53]采用了动态选择抽样策略(Iterative selective sampling, SeS)解决了模型训练过程中正样本过多、数据集不平衡的问题,将训练精力动态地集中在训练难度较大的样本上,从而加快模型训练任务,提高整体性能;在数据集不平衡问题上,Dai等^[54]提出基于临床报告和影像特征的多模态信息整合多尺度卷积神经网络

表3 病变区域检测相关研究
Table 3 Related works on lesion detection

| 相关研究 | 方法 | 数据集 | 提取特征 | 性能 |
|------------------------------------|---------------------------------------|--------------------------|----------------|--|
| Shan等 ^[51] | Patches + 堆叠稀疏自动编码(SSAE) + 迁移学习 | DIARETDB | MA | 敏感性: 91.6%, F-score: 91.3%, 准确性: 91.38% |
| Budak等 ^[55] | 深度卷积神经网络(DCNN) | 在线挑战数据集(ROC) | MA | 比赛分数为0.221, 高于其他方法 |
| Dai等 ^[54] | Alex-Net 框架为基础的MS-CNN模型 | 当地医院收集数据, DIARETDB1 | MA | 准确率: 96.1% |
| Orlando等 ^[56] | CNN + 手工工程特征 + 随机森林(Random forest)分类器 | MESSIDOR E-Ophtha | MA, HE | AUC: CNN: 0.7912, 手工工程: 0.7325, CNN+手工工程: 0.8932 AUC: CNN: 0.8374, 手工工程: 0.8812, CNN+手工工程: 0.9031 |
| van Grinsven等 ^[53] | 动态选择抽样策略(SeS, NSeS) + 10层CNN | Kaggle MESSIDOR | HE | 敏感性: 84.8%, 特异性: 90.4%, AUC: 91.7%, 敏感性: 93.1%, 特异性: 91.5%, AUC: 97.9% |
| Prentas等 ^[57] | DNN | DRiDB | EX, SE | 敏感性: 78%, F-score: 78% |
| Otalora等 ^[58] | 基于LeNet网络 + EGL的主动学习策略 + 迁移学习 | E-Ophtha | EX, SE | 敏感性: 99.8%, 特异性: 99.6%, 准确性: 99.6% |
| Abbasi-Sureshjani等 ^[52] | ResNet | DIARETDB1, DR2, E-Ophtha | EX, SE | AUC 分别为: 96.5%, 97.2%, 99.4% |
| Badar等 ^[40] | SegNet 模型为基础的Auto-encoder 网络 | Messidor | MA, HE, EX | 精确度: 99.24% (EX), 97.86% (HM), 88.65% (MA) |
| ISBI 韩国 VRT 团队 | DeepLab + U-Net | DRiDB | MA, HE, SE, EX | F-score 分别为: 0.4951, 0.6804, 0.6995, 0.7127 |
| ISBI 中国平安科技 Patech 团队 | DenseNet + DeepLab V3 | DRiDB | MA, HE, EX | F-score 分别为: 0.474, 0.649, 0.885 |
| ISBI 中国科大讯飞 | U-Net + 注意力机制 + DeepLabV3 + | DRiDB | MA, HE, SE, EX | F-score 分别为: 0.5017, 0.5588, 0.6588, 0.8741 |
| Tan等 ^[59] | DCNN | CLEOPATRA | MA, HM, HE, SE | 敏感性: 87.58%, 特异性: 98.73% |

(Multi-scale convolutional neural network, MS-CNN) 策略, 可以克服数据集不平衡、成像条件多样化、弥补视觉特征提取不充分等问题, 可尝试将该方法扩展到眼底影像多种病灶区域的检测, 在医学领域研究多模态方面具有战略性意义; Otálora 等^[58] 采用主动学习策略, 减轻了训练模型中的标注工作, 具有一定的先进性, 但当未标记数据样本数量较大时, 该方法存在计算缺陷, 需要进一步改进; 2018 年 ISBI 中多个队伍用到的 DeepLab 系列模型, 提高了特征图像分辨率、网络的运行速率和健壮性等, 其中中国科大讯飞采用注意力机制, 将注意力集中到影像的特定像素上, 从而减少了任务的复杂度, 具有重要的研究价值. Badar 等^[40] 采用基于 SegNet 的端到端模型同时对多类视网膜病变进行分割, 实现病灶边界的精确定位等等. 深度学习技术在病灶区域检测方面有越来越突出的表现.

以下对这些研究进行详细介绍:

微型动脉瘤和出血特征检测^[60]: 红色的病灶是 DR 最早期可察觉的症状, 这是一个通用的术语, 即包括微型动脉瘤 (MAs) 和出血 (HEs). 精确检测红色病灶对检测早期糖尿病视网膜病变具有重大的意义, 但由于病灶体积小且缺乏对比度, 所以该任务既繁琐又耗时, 非常具有挑战性.

针对以上问题, 2016 年 Shan 等^[51] 提出了基于堆叠稀疏自动编码 (Stacked sparse auto encode, SSAE) 的 MA 检测方法, 从原始眼底影像中生成 Patches, 并在每个 Patch 中自动提取特征, 使用 Softmax 分类器将每个 Patch 分类为 MA 或非 MA. 该方法通过迁移学习在 DIARETDB 数据集进行了训练和测试.

2017 年 Budak 等^[55] 提出一种基于彩色眼底影像的 MAs 检测系统, 该方法由三个阶段组成: 图像预处理; 通过检测 MA 潜在位置, 得到一系列候选 MAs; 最后利用带有增强样本学习策略的深度卷积神经网络 (Deep convolutional neural network, DCNN) 判断 MA 的真实位置, 并在视网膜病变在线挑战数据集上做了大量的实验来验证该模型.

2018 年 Dai 等^[54] 提出了 Multi-sieving CNN (MS-CNN). 该系统通过文本挖掘技术从临床文本报告中提取专家知识, 结合对应的眼底影像, 共同训练弱影像数据到文本数据的映射模型, 对眼底影像的病灶进行粗分割; 然后基于 Alex-Net 模型得到最终分割结果, 该系统在当地医院收集的数据集以及公共 DIARETDB1 数据集上得到验证.

2018 年 10 月 Orlando 等^[56] 提出了一种基于深度学习和领域知识相结合的红色病灶检测新方法, 并将红色病灶的严重程度分为 R0 和 R1 两个等级,

以准确的检测 MA, 并区分更早的病变阶段. 通过深度学习方法结合特征工程共同学习病灶特征, 然后基于随机森林分类器对特征向量进行分类, 识别阳性病灶, 最终在 MESSIDOR 和 E-Ophtha 数据集上进行了验证, 发现将两种方法结合在一起可以显著提高模型性能.

2016 年 van Grinsven 等^[53] 提出了一种动态选择抽样策略, 在模型训练过程中对负采样中的每个像素分配一个动态权重, 表示其附带的信息水平, 每个训练阶段之后, 每个负采样像素的权重都会更新, 重复这个过程, 直到达到预设标准. 最后使用训练好的 CNN 模型对测试图像中的每个像素进行分类, 得到测试图像的像素概率图, 该方法解决了模型训练过程中正样本过多的问题, 并在 Kaggle 和 MESSIDOR 平台上测试了他们的系统.

渗出物是威胁视觉糖尿病视网膜病变和糖尿病黄斑水肿的早期症状之一. 因此, 要防止威胁视力的视网膜病变的发生, 早期的渗出物鉴定是至关重要的, 渗出物分割在评估病变的严重程度中也起着重要的作用^[49].

2016 年 Prentas 等^[57] 提出了一种基于监督深度神经网络的渗出物检测系统. 该系统通过深度卷积神经网络生成渗出物检测概率图, 并检测视盘、血管的位置信息, 通过血管位置、视盘位置以及明亮边界等信息来提高渗出物检测性能. 最后, 使用 DRiDB 数据集分析了这种扩展方法.

2017 年 Otálora 等^[58] 提出了一种基于 EGL (Expected gradient length) 的 Label-efficient CNN 模型, 通过主动学习策略来选择信息量最大的 Patches 和图像, 解决了标注数据样本匮乏的问题. 并且模型比通常的随机梯度下降 (Stochastic gradient descent, SGD) 策略收敛更快, 并通过生产掩码, 预测和分割病灶的区域. 该模型基于 LeNet 网络^[61] 采用迁移学习方法进行训练, 利用公开的数据集 E-Ophtha 测试了该算法的有效性.

2017 年 Abbasi-Sureshjani 等^[52] 提出了一种基于 ResNet 模型的全自动渗出分割方法, 与 Shan 等^[51] 提出的方法相似, 在 Patches 上进行训练, 该方法针对训练过程中存在反射和其他明亮病灶造成的大量假阳性问题, 提出了选择具有较高优先级的易误导的常规 Patches 方法, 使网络学习自动调整参数, 同时避免了网络对这些样本的过度拟合. 该方法在三个公共数据集 DIARETDB1、DR2 和 E-Ophtha-EX 分别进行网络的测试评估.

上文所述方法只关注一种病理特征的识别, 而 DR 分级要求同时检测多种病灶特征, 越来越多的算法趋向于视网膜影像中多种病理同时分割的研究, 以对视网膜的病变程度进行精确诊断.

2018年 Badar 等^[40] 基于 SegNet 模型, 对视网膜病理语义像素进行描述, 得到高维复杂的特征图, 并为每个像素分配特定类别标签, 如渗出物、出血物或棉毛斑点, 得到像素级分割结果. 该算法在公开的 Messidor 数据集上进行了评估.

2018年4月 ISBI 糖尿病视网膜病变—分割和分级的挑战中, 获得分割挑战赛第1名的韩国 VRT 团队, 通过 DeepLab 思想改进 U-Net 分割模型, 将 U-Net 模型的最大池化层替换成步长为1的 3×3 Atrous convolution, 进行密集特征提取, 实现语义分割; 排名第2的中国平安科技 Patech 团队以 DenseNet 和 DeepLab V3 网络为基础, 实现了病灶特征自动分割; 排名第3的中国科大讯飞, 以 U-Net 和 DeepLabV3+网络为基础, 融合了 Atrous convolution 操作、Self-attention 机制^[62] 的优势, 在不损失原始影像信息的前提下, 快速定位出可疑病灶区域并给出分割边界.

2017年10月 Tan 等^[59] 通过10层深度神经网络模型, 自动分割眼底影像中渗出物、小动脉瘤和出血区域, 对于卷积层和中间全连接层, 采用 leaky Relu 激活函数, 最后一层全连通层采用 softmax 激活函数, 并对眼底影像进行归一化处理^[63], 达到校正光照不均匀、增强局部对比度的目的. 该模型通过 CLEOPATRA 数据集进行训练和测试.

5 糖尿病视网膜病变等级分类

本节介绍了几种常用的深度学习分类模型, 概述了模型改进及发展的方向, 促使 DR 等级分类算法性能的不断改进; 随后对 DR 病变等级分类的相关研究进行综述, 了解最新研究趋势并对其优缺点进行总结, 具体如图3所示.

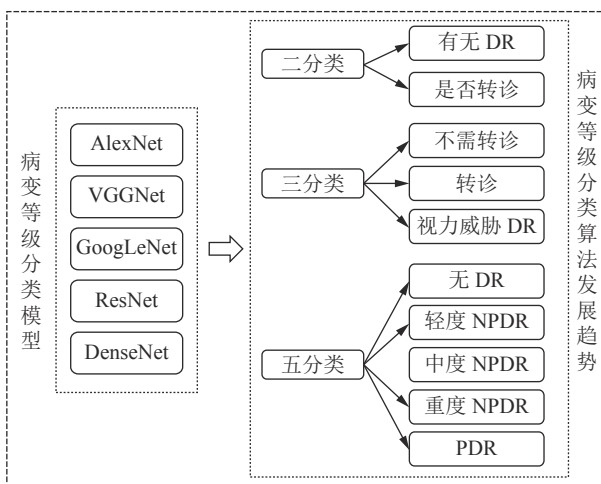


图3 糖尿病视网膜病变等级分类

Fig.3 Classification of diabetic retinopathy

5.1 病变等级分类模型

从2012年的 AlexNet^[64], 2014年的 VGGNet^[38] 和2015年 GoogLeNet^[65] 到2016年的 ResNet^[39], 再到2017年 DenseNet^[66], 分类模型的规模和深度急剧增加, 识别错误率快速降低, 但计算量却并未显著增加. 下面简单介绍几个在糖尿病视网膜研究中常用的分类模型.

1) GoogLeNet 模型

2015年, 谷歌提出了 GoogLeNet^[65] 模型, 在其性能进行提升及优化运算量后进一步提出了 InceptionV2^[67] 及 InceptionV3^[68] 模型, 该模型被相关研究人员广泛应用在 DR 诊断中, 比如2018年10月中山大学团队^[16] 以 InceptionV3 模型为基础实现了对威胁视力的视网膜病变诊断, 以及2016年11月谷歌团队^[69] 同样以 InceptionV3 模型为基础实现了对需要转诊的视网膜病变的诊断等研究.

GoogLeNet 最大的特点就是使用了 Inception 模块, 其目的是设计一种具有局部拓扑结构的网络, 对输入图像并行地执行多个卷积运算或池化操作, 并将所有输出结果拼接为一个非常深的特征图. 随后, 通过2个连续的 3×3 卷积层 (stride = 1) 组成的小网络代替单个的 5×5 卷积层减小计算量, 并提出了著名的批量归一化 (Batch normalization, BN) 思想, 保持感受野范围的同时又减少了参数量; InceptionV3 模型将 $n \times n$ 卷积拆分为 $1 \times n$ 卷积和 $n \times 1$ 卷积, 进一步减少了计算量. 目前很多医学影像检测中使用 InceptionV3 模型进行迁移学习, 以解决医学领域数据匮乏问题.

2) ResNet 模型

2016年, 文献 [39] 提出了深度残差网络 (ResNet)^[39], 将网络深度提升到上百层. ResNet 最大的特点是解决了反向传播过程中的梯度消失问题, 因此它可以训练非常深的网络, 而不用像 GoogLeNet 需要在中间添加分类网络以提供额外的梯度.

ResNet 引入了残差连接, 在每一个残差模块上, 残差连接会将该模块的输入与输出直接相加, 因此在反向传播中, 根据残差连接传递的梯度就可以不经过残差模块内部的多个卷积层, 为前一层保留足够的梯度信息. 在 DR 检测中很多模型通过 ResNet 网络来提高系统的健壮性.

3) DenseNet 模型

2017年 Huang 等^[70] 提出了 DenseNet 网络, 提升了网络层级间信息流与梯度流的效率, 提高了参数利用率. DenseNet 将不同路径组合特征图按深度相互拼接在一起, 使每一层的输出都作为之后所有层的输入, 最后一层将拼接之前所有网络层输出

的特征图. 这种结构确保了每一层能从损失函数直接访问到梯度, 因此可以训练非常深的网络, 在近两年糖尿病视网膜病变分级挑战赛以及研讨会上, 很多研究者采用了 DenseNet 模型的思想, 并取得了突出的效果.

5.2 病变等级分类的相关研究

近年来, 越来越多的研究团队将深度学习算法应用于糖尿病视网膜病变等级分类研究, 如表 4 所示.

最初研究主要集中在有无 DR 或者是否需要转诊两种情况的分类, 其中需转诊的糖尿病视网膜病变 (Referable diabetic retinopathy, RDR) 包括中度 DR, 重度非增值性 DR, PDR, 以及更严重的 DR. 虽然取得了较好的效果, 但是没有提供威胁视

力的糖尿病视网膜病变 VTDR (Vision-threatening diabetic retinopathy) (包括 PDR (Proliferative Diabetic Retinopathy) 及 DME (Diabetic macular edema)) 的可比数据^[71], 这种比较严重的病例通常需要紧急转诊和临床护理, 理想情况下不应该被任何筛查程序 (无论是人还是软件) 所忽视. 由于国际 DR 诊断标准全球公认, 越来越多的研究者趋向于按照此标准, 将 DR 诊断结果分为 R0 (无 DR)、R1 (轻度 NPDR)、R2 (中度 NPDR)、R3 (重度 NPDR)、R4 (PDR) 五类, 以辅助医生制定更合理的治疗方案.

通过对以上研究的回顾表明, DR 智能诊断系统的研究, 向着多分类、稳健性、实用性的方向发展. 2016 年谷歌^[69] 在《美国医学会杂志》中报道了基于深度学习的 DR 检测方法, 该算法能够在视网膜造

表 4 病变等级分类相关研究
Table 4 Related works on classification of diabetic retinopathy

| 相关研究 | 应用 | 方法 | 数据集 | 性能 |
|------------------------------|----------------------|---|---|---|
| 谷歌 Gulshan 等 ^[69] | 诊断 RDR | InceptionV3 框架, 端到端分类 | EyePACS-1 Messidor-2 | 特异性: 93.4%, 敏感性: 97.5% 特异性: 93.9%, 敏感性: 96.1% |
| Li 等 ^[72] | 诊断有无 DR | VggNet, GoogLeNet, Vgg-s 等进行迁移学习 | DR1 MESSIDOR | Vgg-s 性能最好: 敏感性: 97.11%, 特异性: 86.03%, 准确度: 92.01%, AUC: 0.9834 |
| ElTanboly 等 ^[73] | 诊断 RDR | Stacked non-negativityconstraint autoencoder (SNCAE) | 医院获取 OCT 图像 | 准确度: 96% |
| Gargeya 等 ^[74] | 诊断 RDR | Data-driven DNN ResNet, Second-level gradientboosting classifier | EyePACS MESSIDOR | AUC: 0.97 AUC: 0.94 |
| Abràmoff 等 ^[75] | 诊断 RDR, VTDR | DCNN | Messidor-2 | RDR 敏感性: 96.8%, 特异性: 87.0%, AUC: 0.980 从 RDR 分类出 VTDR: 敏感性: 100%, 特异性: 91%, AUC: 0.989 |
| Ting 等 ^[76] | 诊断 RDR, VTDR | VGG-19 | 6 个不同国家招募了 10 组数据集 | RDR 敏感性: 90.5%, 特异性: 91.6%, AUC: 0.936 从 RDR 分类出 VTDR: 敏感性: 100%, 特异性: 91.1%, AUC: 0.958 |
| 中山大学 ^[16] | 诊断 RDR, VTDR | InceptionV3 | 中国人彩色眼底图片 多种族彩色眼底图像 | 敏感性: 97.0%, 特异性: 91.4% 敏感性: 92.5%, 特异性: 98.5% |
| Abràmoff 等 ^[77] | 在初级保健诊所诊断 DR, 进行实际应用 | CNN | 在初级保健诊所招收了 900 名受试者, 男性占 47.5%; 其中包括: 西班牙裔 16.1%, 非裔美国人 28.6% | 敏感性: 87.2%, 特异性: 90.7%, 显像率: 96.1% |
| Wang 等 ^[78] | 划分五类等级 | DenseNet, Boosting tree algorithm | Kaggle | 精确度: R0: 0.92, R1: 0.70, R2: 0.64, R3: 0.67, R4: 0.69 |
| Zhou 等 ^[79] | 划分五类等级 | Multi-Cell Multi-Task CNN | Kaggle | Kappa: 0.841 |
| Doshi 等 ^[80] | 划分五类等级 | 5层 CNN 网络 | EyePACS | Kappa: 0.386 |
| IBM | 划分五类等级 | DCNN | EyePAC | 准确度: 86% |

影中检测到 DR 的存在迹象,吸引了越来越多的研究者展开研究;随后 Abràmoff 等^[75]通过实验证明了基于深度学习的 DR 检测算法具有更高精度,优于未使用深度学习的 DR 检测算法;Li 等^[72]采用迁移学习的方法,从其他相关任务的大数据集中学习知识,并在目标数据集中实现高精度的分类效果,解决了眼底影像数据匮乏的问题;为了更好地诊断轻微 DR,防治严重病情的发生,ElTanboly 等^[73]和 Gargeya 等^[74]分别通过 SNCAE 和 ResNet 模型分类出患有轻微 DR 的患者,以进行早期治疗,其中,ElTanboly 等^[73]使用 OCT 影像对不同 OCT 切片上的眼部细节进行检查,具有更高的准确度;ISBI 会议中文献^[76]采用了 DenseNet 模型,并对每幅图像进行 50 次随机增强,以提取稳定精确的特征,提高分类精度;因为一些小的病理组织只有通过高分辨率视网膜影像才能被发现,但直接通过深度神经网络模型训练高分辨率影像,时间与计算量成本较大;Zhou 等^[79]通过一种多单元多任务结构模型,缓解了以上问题,并且该模型可作为通用架构,很容易与其他深度神经网络模型集成,具有广泛的适用性和重要的研究价值。

尽管以上研究均已在公开可用的数据集上进行验证,展示了 DR 智能诊断系统的优越性,但是并没有在临床 DR 诊断筛选方案中进行测试,未解决不同人种和不同设备的成像差异,针对此问题,Ting 等^[76]和中山大学^[16]等多项研究,通过收集不同种族、不同国家的眼底影像验证 DR 诊断算法的临床可用性,以及由 Abràmoff 等^[77]开发的深度学习系统,在社区初级保健诊所进行测试,取得了较好的效果,并获得了美国食品和药物管理局 (Food and Drug Administration, FDA) 的批准。

以下对这些研究方法进行详细介绍。

2016 年, Gulshan 等^[69]利用深度学习技术对需要转诊的 DR 影像 (RDR) 进行分类。首先,根据国际 DR 诊断标准对来自 EyePACS 和印度三家眼科医院大型数据集 ($n = 128175$) 进行标注,然后,通过以 InceptionV3 为基础的深度学习模型进行训练,并利用 EyePACS-1 数据集及 Messidor-2 数据集测试该算法的性能。

2017 年 10 月 Li 等^[72]采用迁移学习对是否存在糖尿病视网膜病变进行检测,将预训练 CNN 模型 (包括 Vgg-s、VggNet、GoogLeNet 及其变体模型) 进行微调,用于眼底影像特征提取,通过提取特征训练支持向量机实现分类,并在 DR1 和 MESSIDOR 数据库进行验证,实现较好的分类结果。

早期发现视网膜病变是挽救患者视力的关键,

为了更好地辨别轻微 DR 的情况以进行早期治疗,2018 年 ElTanboly 等^[73]在 ISBI 会议上提出一种利用 OCT 影像对 DR 进行检测和分类的新型计算机辅助诊断系统 (CAD),首先通过深度学习模型分割 12 层视网膜层。然后基于堆叠非负性约束自编码器 (SNCAE) 模型,根据视网膜层的三个鉴别特征 (反射率、曲率和厚度),训练两阶段深度融合分类网络,第 1 阶段将患者分类为正常组或 DR 组,第 2 阶段将 DR 组分类为早期组或中度组,从而精确的诊断早期视网膜病变。为了同样的目的,2017 年 Gargeya 等^[74]提出基于 ResNet 模型进行 DR 智能诊断,从最后一层全局平均池化层提取模型学习到的 1024 个特征,构造可视化热力图,并训练一个基于树二分类模型,对 DR 和健康影像分类,区分出轻度 DR,并在 MESSIDOR-2 等数据集上进行验证。

2016 年 Abràmoff 等^[75]证明了基于深度学习增强算法与未采用深度学习的爱荷华州检测项目算法相比,在 NRDR, RDR 和 VTDR 识别方面有了很大的提高,通过 Messidor-2 数据集进行训练,如果 VTDR 索引高于或等于阈值,则返回 VTDR 的正输出,如果 VTDR 索引低于此阈值,则将对 RDR 索引进行阈值化,如果 RDR 索引高于或等于后一个阈值,则返回 RDR 的正输出,如果低于后一个阈值,则返回 NRDR。通过对比,该算法 RDR 的敏感性与 IDP 算法敏感性无统计学差异,但特异度明显优于 IDP,因此,更适合于潜在的糖尿病患者进行筛查,以尽量减少假阳性读数。

为了提供 DR 智能诊断系统的临床可用性,2017 年 Ting 等^[76]提出了一种基于 VGG-19 开发的深度学习系统,在新加坡综合糖尿病视网膜病变检测项目中测试了 5 年,并从 6 个不同国家收集了 10 组数据,对系统进行验证。2018 年 10 月中山大学^[16]提出了一种基于 Inception V3 的深度学习算法 (DLA),该算法通过使用一组 106244 幅眼底影像数据进行训练测试,其中包括 71043 幅中国人彩色眼底影像和 35201 幅多种族彩色眼底影像。2018 年 1 月 Abràmoff 等^[77]提出基于深度学习的 DR 诊断系统,并在初级诊所 (包括不同种族、不同年龄人群的眼底影像) 对 DR 诊断进行测试,并成为第一个 FDA 授权的 DR 智能诊断系统。这表明人工智能有能力将专业诊断引入初级保健诊所。

2018 年 4 月 ISBI 糖尿病视网膜病变—分割和分级的挑战中,文献^[78]首先对输入眼底影像进行预处理,突出 DR 病灶信息,然后通过 121 层的 DenseNet 模型提取 DR 病灶特征,并进行 50 次随

机增强,从最后一层全连接层得到 50 个输出,计算输出值的均值向量和标准差向量作为特征值,通过增强树算法进行预测,并在 Kaggle 数据集进行测试,验证了该方法在对视网膜影像进行分级时比仅使用 DenseNet 模型具有更好的性能。

为了提高深度学习的运行速度,2016 年 12 月 Doshi 等^[80]通过 GPU 加速深度卷积神经网络训练过程,实现了 DR 智能诊断,并根据病情严重程度将视网膜影像划分为 5 个等级,并在 EyePACs 数据集上进行验证。2018 年 Zhou 等^[79]将高分辨率眼底影像通过深度神经网络进行训练,通过多单元结构节省了训练时间。此外,考虑到 DR 不同病变阶段实际上是渐进的,意味着不同阶段的标签是相关的,为了考虑不同病变等级影像之间的关系,提出一种多任务学习策略,该策略通过分类和回归来预测标签,并在 Kaggle 数据集验证了该方法的有效性。

2017 年 IBM,通过基于深度学习的算法对来自 EyePACs 的 35000 多幅眼底影像识别微动脉瘤、出血和渗出物等病变区域,评估糖尿病视网膜病变的严重程度,并能在 20 s 内给出诊断结果,临床医生可以利用该系统来更好地了解疾病的发生发展,并制定有效的治疗方案。

6 基于 OCT 影像的眼部疾病诊断

视网膜 OCT 影像作为一种非侵入性和非接触性的成像方式,广泛应用于眼科疾病诊断,如 DR (包括并发症糖尿病性黄斑水肿 (DME))、年龄相关性黄斑变性 (Age-related macular degeneration,

AMD)、青光眼等,可通过分割视网膜内囊肿 (Intraretinal cyst, IRC)、渗出物、视网膜层、色素上皮脱离 (Pigment retinal epithelium, PRE) 以及视网膜下积液 (Subretinal fluid, SRF) 等病灶特征,并进行量化分析,可辅助医生同时对多种眼部疾病做出诊断。因此,基于深度学习自动识别 OCT 影像具有重要意义,引起了广泛的研究热潮。基于 OCT 影像的眼部疾病研究如表 5 所示。

通过对以上研究的回顾表明,基于深度学习的眼部疾病检测已经广泛应用于 OCT 分割、分类以及转诊等方面,并取得了突破性的成果。在 OCT 影像分割方面,深度学习网络可成功分割 IRC、SRF 以及视网膜层紊乱等病灶, Sandhu 等^[84]通过对 OCT 分层,并对每层视网膜层进行量化分析,实现 DR 严重程度进行分类。Hassan 等^[81]提出了一种稳健的分割模型,通过相干张量从 OCT 影像中提取具有不同病理综合征的视网膜层,即使发生细微变化,也能对病症自动提取; Vahadane 等^[82]通过对硬渗出物和囊肿区域进行分割,预测 DME 的存在。在实现眼部疾病转诊方面, Kermany 等^[15]和 Li 等^[83]均采用迁移学习的方法,对 OCT 影像进行诊断转诊,并且 Kermany 等^[15]通过遮挡实验检测模型感兴趣区域 (ROI) 作为判断的依据,提高了模型的可解释性;谷歌 DeepMind 团队^[18]通过两阶段深度神经网络,不仅实现了转诊,而且将 OCT 影像分割成不同的眼部组织,缓解了深度学习模型可解释性问题,并且通过分割模型可以对不同供应商的 OCT 设备进行泛化,而不需要严格的再训练过程,在其他医学影像检测方面具有广泛的借鉴价值。以下对

表 5 基于视网膜 OCT 影像的眼部疾病诊断相关研究
Table 5 Studies on diagnosis of ocular diseases based on retinal OCT images

| 相关研究 | 应用 | 方法 | 数据集 | 性能 |
|-----------------------------|----------------------|---------------------------------|---|---|
| Sandhu 等 ^[84] | 对 DR 患者早期诊断 | 基于融合形状、强度和空间信息的联合模型,两阶段深度融合分类网络 | 路易斯维尔大学 (University of Louisville) 接受常规筛查和 (或) 监测检查的 II 型糖尿病患者 OCT 数据 | 准确率平均为 94% |
| Hassan 等 ^[81] | 从 OCT 扫描中分割出 8 个视网膜层 | 深度卷积神经网络,基于结构张力的分割框架 (CNN-STSF) | 取自不同公共可用数据集和当地武装部队眼科研究所 (AFIO) 数据集的超过 3.9 万幅视网膜 OCT 影像 | 准确性: 93.75% |
| Vahadane 等 ^[82] | 分割硬渗出物和囊肿区域检测 DME | 图像处理,深度学习,基于规则方法 | 1827 幅 OCT 影像 | Precision: 96.43%, Recall: 89.45%, F1-score: 92.81% |
| Kermany 等 ^[15] | 疾病分类转诊 | InceptionV3+ 迁移学习,分类网络 | 108312 幅 OCT 训练影像 (37206 幅脉络膜新生血管, 11349 幅糖尿病黄斑水肿, 8617 幅玻璃膜疣, 51140 幅正常), 1000 幅 OCT 测试影像 (每个类别 250 幅) | 准确率: 96.6%, 敏感性: 97.8%, 特异性: 97.4%, AUC: 99.9% |
| Li 等 ^[83] | 疾病分类转诊 | VGG-16 + 迁移学习,分类网络 | 医院获取 109312 幅 OCT 影像 (37456 幅脉络膜新生血管, 11599 幅糖尿病黄斑水肿, 8867 幅玻璃膜疣, 51390 幅正常) | 准确率: 98.6%, 敏感性: 97.8%, 特异性: 99.4%, AUC: 100% |
| DeepMind 团队 ^[18] | 分割疾病特征分类转诊 | 3DU-NET 分割网络, CNN 分类网络 | 摩尔菲尔兹眼科医院提供 14 884 幅 OCT 影像 | 准确度: 94%, AUC: 99.21% |

这些研究进行详细介绍。

2018 年 Sandhu 等^[84] 开发了一种新的无创框架, 利用 OCT 影像对 DR 进行早期诊断. 首先, 采用融合形态学、空间和强度信息的联合模型对 12 层不同的视网膜层进行分割; 其次采用两阶段深度融合分类网络, 根据所有视网膜层的曲率、反射率和厚度三个鉴别特征, 将受试者分为正常、亚临床期 DR 或轻度/中度 DR, 其中亚临床 DR 定义为临床眼底检查为阴性。

2018 年 Hassan 等^[81] 提出了一种深度卷积神经网络和基于结构张力的分割框架 (CNN-STSF), 用于从正常和病变 OCT 影像中自动分割 8 层视网膜层. 首先, 在视网膜影像中提取多个候选视网膜层 Patches, 并计算相干张量; 然后, 利用基于云的深度卷积神经网络模型, 对 1 200 个视网膜层 Patches 进行训练、分类, 分割精度达到 0.9375, 可辅助医生根据视网膜层厚度、曲率等特征对 DR、AMD 等疾病做出诊断。

2018 ISBI 挑战赛中, Vahadane 等^[82] 提出了一个两阶段的框架来分割硬渗出物和囊肿区域, 检测 DME. 第 1 步使用图像处理技术检测硬渗出物和囊肿区域的候选 Patches; 第 2 步是使用深度卷积神经网络预测这些候选 Patches 的标签. 在最后的排序步骤中, 聚合 CNN 模型的置信度, 并使用基于规则的方法来预测 DME 的存在, 较其他参赛方法取得了较好的效果。

2018 年 Kermany 等^[15], 2019 年 Li 等^[83] 均采用迁移学习的方式, 分别将训练好的 InceptionV3 网络、VGG-16 网络进行微调, 用于 OCT 影像检测, 可诊断脉络膜新生血管 (Choroidal neovascularization, CNV), DME, 玻璃膜疣 (Drusen), 正常 (Normal) 等 4 种疾病, 并提供转诊建议 (包括紧急转诊 (CNV、DME)、常规转诊 (Drusen)、仅观察 (Normal)), 并且 Kermany 通过热力图的方式提高了模型的可解释性。

谷歌公司 DeepMind 团队^[18] 通过在两阶段深度神经网络实现 OCT 影像自动分割及转诊, 第 1 个网络是 3D U-Net 分割网络, 能分割出多种疾病特征, 例如黄斑部视网膜水肿 (Macular retinal edema, MRE)、脉络膜新生血管 (CNV)、玻璃膜疣 (Drusen)、视网膜前膜 (ERM) 以及 DME 等; 第 2 个网络为 CNN 分类网络, 通过分析病灶的位置、大小及数量信息, 准确检测 50 余种危害视力的眼部疾病, 并为患者提供转诊建议 (包括紧急转诊、半紧急转诊、常规转诊、仅观察), 准确度可媲美顶级专家医生。

7 模型评估

本小节主要介绍了通常用于评估 DR 检测算法的性能指标. 常用的度量标准有敏感度 (Sensitivity)、特异度、准确性、精确性、F-score、ROC 曲线以及一致性评估标准 Kappa 值^[85] 等。

通常对于一个样本的预测类别和真实类别, 有四种可能组合分别是:

- 1) TP (True positive): 预测和真实类别都为阳性;
- 2) TN (True negative): 预测和真实类别都为阴性;
- 3) FP (False positive): 预测为阳性, 真实类别为阴性;
- 4) FN (False negative): 预测为阴性, 真实类别为阳性的。

敏感度 (Sensitivity)、特异度、准确性、精确性都可以通过 TP、TN、FP、FN 来表示^[86]。

敏感度 (Sensitivity, 也称召回率 (Recall)) 用来表示实际为阳性的样本中判断为阳性的比例, 计算式为

$$Sen = \frac{TP}{TP + FN} \quad (1)$$

敏感度的值越大, 说明“漏检” (FN) 越少。

特异度 (Specificity) 是指实际为阴性的样本中, 判断为阴性的比例, 计算式为

$$Spe = \frac{TN}{TN + FP} \quad (2)$$

特异度的值越大, 说明“误检” (FP) 越少^[87]。

对于任何分类器而言, 都需要在敏感度及特异度之间进行取舍. 这个取舍可以通过接收器操作特征 (Receiver operating characteristic, ROC) 曲线来表示^[85]. AUC (Area under curve) 即 ROC 特征曲线下面积, 常用来衡量一个模型的总体准确情况, 取值范围在 0 到 1 之间, AUC 越接近 1, 性能越好。

正确率 (Accuracy) 表示对于给定的测试数据集, 正确分类的样本数与总样本数的比^[86], 计算式为

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

精确率 (Precision) 表示预测阳性中准确的占比, 计算式为

$$Pre = \frac{TP}{TP + FP} \quad (4)$$

F-score 将精确率和敏感度结合如下:

$$F = 2 \times \frac{Pre \times Sen}{Pre + Sen} \quad (5)$$

一致性评估的评估标准为 Kappa 值^[87]. 目前, 针对 DR 检测最大的公共数据集^[88], 即 Kaggle 的 DR 检测竞赛 (Kaggle diabetic retinopathy detection competition) 数据^[89], 该竞赛中, 模型的优化目标为最小的平方加权 Kappa 值, 表示模型诊断结果与专家诊断结果的一致性^[90].

Kappa 计算式为

$$\kappa = \frac{P_0 - P_c}{1 - P_c} \quad (6)$$

其中, P_0 为观测一致性, 表示两种检测结果一致的百分比; P_c 为期望一致性, 表示两种检测结果预期相同的概率^[89].

Kappa 值计算的结果为 $-1 \sim 1$, 但通常落在 $0 \sim 1$ 间, 可分为五组来表示不同等级的吻合度: $0.0 \sim 0.20$ 极低的吻合度、 $0.21 \sim 0.40$ 一般的吻合度、 $0.41 \sim 0.60$ 中等的吻合度、 $0.61 \sim 0.80$ 高度的吻合度和 $0.81 \sim 1$ 几乎完全吻合^[90]. Kappa 值在临床上常用来确定不同检验或诊断方法的一致性, 因此常作为重要的模型一致性评估标准.

除 Kappa 值之外, 一些论文引入权重矩阵、惩罚模型, 对不同的程度错误赋以不同的权重, 例如惩罚假阴性的严重错误分类. 每个专家的决策都会经过权重矩阵的加权. 最后计算了每个分类结果的平均误差.

8 挑战

深度学习在糖尿病视网膜疾病诊断方面有着良好的应用前景, 但仍面临许多挑战:

1) 高质量标注数据的获取难度大^[91]. 深度学习的参数具有复杂性, 因此需要使用足够的带有 Ground-truth 标签的数据进行训练^[92]. 为了获取高质量的标注数据, 研究人员必须聘请多名临床医生对病理图像进行标注, 这一过程可能是耗时且昂贵的^[93], 增加了研究的难度与成本.

2) 隐私问题. 医学数据中往往涉及到患者的隐私, 患者隐私会严重限制医学数据的共享和使用^[94], 即使没有共享数据, 在机密患者数据上训练的算法也可能存在安全风险^[95] 或被攻击者窃取数据^[96]. 如何在进行科学研究的同时, 保证病人的隐私不被泄露, 是研究者应当考虑的重要问题. 目前这个方向上也已经提出了技术解决方案, 例如差异隐私框架下训练深度学习模型^[97].

3) 深度学习的可解释性问题. 深度学习的内部机制尚不完全透明, 这是阻碍深度学习辅

助疾病诊断实际应用的重要因素之一. 如果能探索其诊断的原理, 并验证其有效性, 深度学习诊断糖尿病视网膜病变的进程将被大大推进^[98]. 另外, 2016 年 4 月欧盟通过的关于个人信息使用的新规“一般数据保护条例”^[99] 提出了“解释权”, 即运用机器学习算法的人必须能够解释算法如何做出决定. 因此, 研究人员需要考虑深度学习模型是否可以被充分解释, 来避免潜在的不良后果^[100].

4) 跨设备数据模型有效性. 不同设备的成像特点可能不同, 控制不同数据集的偏差将可以限制统计显著性、防止检测弱效应和限制模型中训练的参数量等^[101]. 而现有的模型往往在特定设备采集的数据中训练而成^[102], 在不同设备所采集的数据集中进行验证时表现一般^[103]. 因此, 保证跨设备数据模型有效性是增加程序可移植性的重要一步.

因此, 深度学习领域和医学界的密切合作至关重要, 通过医工交叉解决以上问题, 不仅要促进深度学习技术的开发和验证, 而且要战略性地将这些技术用于患者护理.

9 结束语

综上所述, 基于深度学习的糖尿病视网膜病变 (DR) 诊断已经成为医学影像分析领域深度学习发展最迅速、诊断精度高、临床试行和智能诊断仪器 FDA 批准进程最快的方向. 在学术界, DR 诊断的敏感性、特异度不断提升, 实验结果显示深度学习模型已经基本达到临床诊断的能力; 在产业界, 已经有相关器械落地^[19], 并朝着便携化的方向发展^[20], 以配合远程医疗的部署.

在病变区域检测方面, 深度学习模型实现了从检测单一病变到支持多种病变综合检测的进步; 在病变等级分类方面, 深度学习模型已经支持国际 DR 分级标准, 实现五等级的划分. 但由于医疗诊断需要明确的依据, 深度学习的可解释性问题使得模型无法满足以上要求, 因而引起学术界的重点研究和研究.

本文旨在为为该领域的研究者与产业界专家提供一定借鉴与参考. 随着技术的进一步发展, 本文所涉及的相关内容也需要进行进一步的补充和完善.

References

- 1 Dhoot D S, Baker K, Saroj N, Vitti R, Berliner A J, Metz C, et al. Baseline factors affecting changes in diabetic retinopathy severity scale score after intravitreal aflibercept or laser for diabetic macular edema: Post hoc analyses from VISTA and VIVID. *Ophthalmology*, 2018, **125**(1): 51-56
- 2 Yang Ling, Shen Xi. Research on correlation between diabetic retinopathy and dry eye. *International Eye Science*, 2018, **18**(4):

- 744-747
(杨玲, 沈玺. 糖尿病性视网膜病变与干眼的相关性研究. 国际眼科杂志, 2018, **18**(4): 744-747)
- 3 Zhang Qiao-Li, Zhao Di, Chi Xue-Bin. Review for deep learning based on medical imaging diagnosis. *Computer Science*, 2017, **44**(S2): 1-7
(张巧丽, 赵地, 迟学斌. 基于深度学习的医学影像诊断综述. 计算机科学, 2017, **44**(S2): 1-7)
- 4 Ciulla T A, Amador A G, Zinman B. Diabetic retinopathy and diabetic macular edema: Pathophysiology, screening, and novel therapies. *Diabetes Care*, 2003, **26**(9): 2653-2664
- 5 Sun Yang, Li Qi-Huan, Tian Si-Jia, Zhang Jie, Wang You-Xin, Ping Zhao, et al. Association of immunoglobulin G N-glycosylation and diabetic retinopathy. *China Preventive Medicine*, 2018, **19**(10): 734-737
(孙扬, 李琪欢, 田思佳, 张杰, 王友信, 平昭, 等. 免疫球蛋白 G N-糖基化与糖尿病视网膜病变的相关性研究. 中国预防医学杂志, 2018, **19**(10): 734-737)
- 6 Palis A G, Golnik K C, Mayorga E P, Filipe H P, Garg P. The international council of ophthalmology 360-degree assessment tool: Development and validation. *Canadian Journal of Ophthalmology*, 2018, **53**(2): 145-49
- 7 Kaur S, Singh D. Early detection and classification of diabetic retinopathy using empirical transform and SVM. In: Proceedings of the 2018 Computational Vision and Bio Inspired Computing. Coimbatore, India: Springer, 2018. 1072-1083
- 8 Wong T Y, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh V C, et al. Guidelines on diabetic eye care: The international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology*, 2018, **125**(10): 1608-1622
- 9 Gass J D M, Sever R J, Sparks D, Goren J. A combined technique of fluorescein funduscopy and angiography of the eye. *Archives of Ophthalmology*, 1967, **78**(4): 455-461
- 10 Hernandez-Matas C, Zabulis X. Super resolution for funduscopy based on 3D image registration. In: Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Chicago, IL, USA: IEEE, 2014. 6332-6338
- 11 Sun Yan-Kui. Medical image processing techniques based on optical coherence tomography and their applications. *Optics and Precision Engineering*, 2014, **22**(4): 1086-1104
(孙延奎. 光学相干层析医学图像处理及其应用. 光学精密工程, 2014, **22**(4): 1086-1104)
- 12 Lin A D, Lee A Y, Zhang Q Q, Rezaei K A, Kinyoun J, Wang R K, et al. Association between OCT-based microangiography perfusion indices and diabetic retinopathy severity. *British Journal of Ophthalmology*, 2017, **101**(7): 960-964
- 13 Eltanboly A H, Palacio A, Shalaby A M, Switala A E, Helmy O, Schaal S, et al. An automated approach for early detection of diabetic retinopathy using SD-OCT images. *Frontiers in Bioscience*, 2018, **10**(1): 197-207
- 14 Muiesan M L, Salvetti M, Paini A, Riviera M, Pintossi C, Bertacchini F, et al. Ocular fundus photography with a smartphone device in acute hypertension. *Journal of Hypertension*, 2017, **35**(8): 1660-1665
- 15 Kermany D S, Goldbaum M, Cai W J, Valentim C C S, Liang H Y, Baxter S L, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 2018, **172**(5): 1122-1131
- 16 Li Z X, Keel S, Liu C, He Y F, Meng W, Scheetz J, et al. An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. *Diabetes Care*, 2018, **41**(12): 2509-2516
- 17 Mookiah M R K, Acharya U R, Chua C K, Lim C M, Ng E Y K, Laude A. Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in Biology & Medicine*, 2013, **43**(12): 2136-2155
- 18 De Fauw J, Ledsam J R, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 2018, **24**(9): 1342-1350
- 19 IDX. Press release: FDA permits marketing of IDx-DR for automated detection of diabetic retinopathy in primary care [Online], available: <https://www.eyediagnostics.net/press-releases/press-release-fda-permits-marketing-of-idx-dr-for-automated-detection-of-diabetic-retinopathy-in-primary-care>, April 12, 2018
- 20 Rajalakshmi R, Subashini R, Anjana R M, Mohan V. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye*, 2018, **32**(6): 1138-1144
- 21 Guo Xiao-Ya. Songyue robot makes an amazing appearance. *Chinese Hospital CEO*, 2018, (14): 28-29
(郭潇雅. 嵩岳机器人惊艳亮相. 中国医院院长, 2018, (14): 28-29)
- 22 Oloumi F, Rangayyan R M, Eells A L. Computer-aided diagnosis of proliferative diabetic retinopathy. In: Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine & Biology Society. San Diego, CA, USA: IEEE, 2012. 1438-1441
- 23 Ting D S W, Cheung G C M, Wong T Y. Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review. *Clinical & Experimental Ophthalmology*, 2016, **44**(4): 260-277
- 24 Porwal P, Pachade S, Kamble R, Kokare M, Deshmukh G, Sahasrabudhe V, et al. Indian diabetic retinopathy image dataset [Online], available: <https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid>, November 13, 2018
- 25 Decenci re E. TeleOphta: Machine learning and image processing methods for teleophthalmology [Online], available: <http://www.adcis.net/en/Download-Third-Party/E-Ophtha.html>, January 10, 2013
- 26 Prenta sic P. Diabetic retinopathy image dataset [Online], available: https://ipg.fer.hr/ipg/resources/image_database, 2013.
- 27 French Ministry. Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology [Online], available: <http://www.adcis.net/en/Download-Third-Party/Messidor.html>, 2014.
- 28 LaTIM laboratory. Messidor 2 [Online], available: <http://latim.univ-brest.fr/indexfce0.html>
- 29 Kauppi T, Kalesnykiene V, Kamarainen J K, Lensu L, Sorri I, Pietil  J, et al. DIARETDB0 — Standard diabetic retinopathy database: Calibration level 0 [Online], available: <http://www.it.lut.fi/project/imageret/diaretdb0/index.html>, May 30, 2007
- 30 Kauppi T, Kalesnykiene V, Kamarainen J K, Lensu L, Sorri I, Raninen A, et al. DIARETDB1-Standard diabetic retinopathy database: Calibration level 1 [Online], available: <http://www.it.lut.fi/project/imageret/diaretdb1/index.html>

- it.lut.fi/project/imageret/diaretddb1/, June 19, 2007
- 31 Rasti R, Rabbani H, Mehridehnavi A, Hajizadeh F. Macular OCT classification using a multi-scale convolutional neural network ensemble. *IEEE Transactions on Medical Imaging*, 2018, **37**(4): 1024–1034
 - 32 Yu Guan-Zhen, Wei Pei-Lian, Chen Ying, Zhu Ming-Hua. Artificial intelligence in pathological diagnosis and assessment of human solid tumor: Application and thinking. *Academic Journal of Second Military Medical University*, 2017, **38**(11): 1349–1354 (于观贞, 魏培莲, 陈颖, 朱明华. 人工智能在肿瘤病理诊断和评估中的应用与思考. 第二军医大学学报, 2017, **38**(11): 1349–1354)
 - 33 Wu bo. AI medical imaging landing actual combat: Data annotation, algorithm method, computing power optimization [Online], available: <http://www.mooc.ai/course/384/learn?lessonid=2165#lesson/2165>, December 21, 2017
 - 34 Karger D R, Oh S, Shah D. Efficient crowdsourcing for multi-class labeling. *ACM SIGMETRICS Performance Evaluation Review*, 2013, **41**(1): 81–92
 - 35 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 3431–3440
 - 36 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(12): 2481–2495
 - 37 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015. 234–241
 - 38 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556, 2014.
 - 39 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016. 770–778
 - 40 Badar M, Shahzad M, Fraz M M. Simultaneous segmentation of multiple retinal pathologies using fully convolutional deep neural network. In: Proceedings of Annual Conference on Medical Image Understanding and Analysis. Southampton, United Kingdom: Springer, 2018. 313–324
 - 41 He K M, Zhang X Y, Ren S Q, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. In: Proceedings of European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014. 346–361
 - 42 Chen L C, Papandreou G, Kokkinos I, Murphy K, Yuille A L. Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiv preprint arXiv: 1412.7062, 2014.
 - 43 Chen L C, Papandreou G, Kokkinos I, Murphy K, Yuille A L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **40**(4): 834–848
 - 44 Lin T Y, Dollár P, Girshick R, He K M, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017. 936–944
 - 45 Chen L C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv: 1706.05587, 2017.
 - 46 Chen L C, Zhu Y K, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. arXiv preprint arXiv: 1802.02611, 2018.
 - 47 Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017. 1800–1807
 - 48 Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, MA, United States: MIT Press, 2014. 3320–3328
 - 49 Wang D X, Cui P, Zhu W W. Deep asymmetric transfer network for unbalanced domain adaptation. In: Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, Louisiana, USA: AAAI, 2018.443–450
 - 50 Dou Q, Ouyang C, Chen C, Chen H, Heng P A. Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss. arXiv preprint arXiv: 1804.10916, 2018.
 - 51 Shan J, Li L. A deep learning method for microaneurysm detection in fundus images. In: Proceedings of the 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). Washington, DC: IEEE, 2016. 357–358
 - 52 Abbasi-Sureshjani S, Dashtbozorg B, ter Haar Romeny B M, Fleuret F. Boosted exudate segmentation in retinal images using residual nets. In: Proceedings of the 2017 Fetal, Infant and Ophthalmic Medical Image Analysis. Québec City, Canada: Springer, 2017. 210–218
 - 53 van Grinsven M J J P, van Ginneken B, Hoyng C B, Theelen T, Sánchez C I. Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1273–1284
 - 54 Dai L, Fang R G, Li H T, Hou X H, Sheng B, Wu Q, et al. Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. *IEEE Transactions on Medical Imaging*, 2018, **37**(5): 1149–1161
 - 55 Budak U, Şengür A, Guo Y H, Akbulut Y. A novel microaneurysms detection approach based on convolutional neural networks with reinforcement sample learning algorithm. *Health Information Science and Systems*, 2017, **5**(1): 14
 - 56 Orlando J I, Prokofyeva E, del Fresno M, Blaschko M B. An ensemble deep learning based approach for red lesion detection in fundus images. *Computer Methods and Programs in Biomedicine*, 2018, **153**: 115–127
 - 57 Prentas P, Lončarić S. Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion. *Computer Methods and Programs in Biomedicine*, 2016, **137**: 281–292
 - 58 Otálora S, Perdomo O, González F, Müller H. Training deep convolutional neural networks with active learning for exudate classification in eye fundus images. In: Proceedings of the 2017

- Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis. Québec City, Canada: Springer, 2017. 146–154.
- 59 Tan J H, Fujita H, Sivaprasad S, Bhandary S V, Raod A K, Chua K C, et al. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Information Sciences*, 2017, 420: 66–76
- 60 Mansour R F. Evolutionary computing enriched computer-aided diagnosis system for diabetic retinopathy: A survey. *IEEE Reviews in Biomedical Engineering*, 2017, 10: 334–349
- 61 LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, **86**(11): 2278–2324
- 62 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, United States: Curran Associates Inc., 2017. 6000–6010
- 63 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 2015 International Conference on Machine Learning. PMLR, 2015: 448–456
- 64 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Red Hook, NY, United States: Curran Associates Inc., 2012. 1097–1105
- 65 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 1–9
- 66 Lu D H, Heisler M, Lee S, Ding G, Sarunic M V, Beg M F. Retinal fluid segmentation and detection in optical coherence tomography images using fully convolutional neural network. arXiv preprint arXiv: 1710.04778, 2017.
- 67 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv: 1502.03167, 2015.
- 68 Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 2818–2826
- 69 Gulshan V, Peng L, Coram M, Stumpe M C, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 2016, **316**(22): 2402–2410
- 70 Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017. 2261–2269
- 71 Wong T Y, Bressler N M. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA*, 2016, **316**(22): 2366–2367
- 72 Li X G, Pang T T, Xiong B, Liu W X, Liang P, Wang T F. Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. In: Proceedings of the 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Shanghai, China: IEEE, 2017. 1–11
- 73 ElTanboly A, Ghazaf M, Khalil A, Shalaby A, Mahmoud A, Switala A, et al. An integrated framework for automatic clinical assessment of diabetic retinopathy grade using spectral domain OCT images. In: Proceedings of the 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington, DC, USA: IEEE, 2018. 1431–1435
- 74 Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*, 2017, **124**(7): 962–969
- 75 Abràmoff M D, Lou Y Y, Erginay A, Clarida W, Amelon R, Folk J C, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investigative Ophthalmology & Visual Science*, 2016, **57**(13): 5200–5206
- 76 Ting D S W, Cheung C Y L, Lim G, Tan G S W, Quang N D, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*, 2017, **318**(22): 2211–2223
- 77 Abràmoff M D, Lavin P T, Birch M, Shah N, Folk J C. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Medicine*, 2018, **1**: 39
- 78 Wang Y, Wang G A, Fan W G, Li J X. A deep learning based pipeline for image grading of diabetic retinopathy. In: Proceedings of the International Conference on Smart Health. Switzerland: Springer, 2018. 240–248
- 79 Zhou K, Gu Z W, Liu W, Luo W X, Cheng J, Gao S H, et al. Multi-cell multi-task convolutional neural networks for diabetic retinopathy grading. In: Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI, USA: IEEE, 2018. 2724–2727
- 80 Doshi D, Shenoy A, Sidhpura D, Gharpure P. Diabetic retinopathy detection using deep convolutional neural networks. In: Proceedings of the 2016 International Conference on Computing, Analytics and Security Trends (CAST). Pune, India: IEEE, 2016. 261–266
- 81 Hassan T, Usman A, Akram M U, Masood M F, Yasin U. Deep learning based automated extraction of intra-retinal layers for analyzing retinal abnormalities. In: Proceedings of the 20th International Conference on e-Health Networking, Applications and Services (Healthcom). Ostrava, Czech Republic: IEEE, 2018. 1–5
- 82 Vahadane A, Joshi A, Madan K, Dastidar T R. Detection of diabetic macular edema in optical coherence tomography scans using patch based deep learning. In: Proceedings of the 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington DC, USA: IEEE, 2018. 1427–1430
- 83 Li F, Chen H, Liu Z, Zhang X D, Wu Z Z. Fully automated detection of retinal disorders by image-based deep learning. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 2019, **257**(3): 495–505
- 84 Sandhu H S, Eltanboly A, Shalaby A, Keynton R S, Schaal S, El-Baz A. Automated diagnosis and grading of diabetic retinopathy using optical coherence tomography. *Investigative Ophthalmology & Visual Science*, 2018, **59**(7): 3155–3160
- 85 Powers D M W. Evaluation: From precision, recall and F-meas-

- ure to ROC, informedness, markedness and correlation. arXiv Preprint arXiv: 2010.16061, 2020.
- 86 Goutte C, Gaussier E. A probabilistic interpretation of precision, recall and F -score, with implication for evaluation. In: Proceedings of the European Conference on Information Retrieval. Santiago de Compostela, Spain: Springer, 2005. 345–359
- 87 Zou K H, O'Malley A J, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 2007, **115**(5): 654–657
- 88 Wu Jiang, Hou Shao-Xin, Jin Meng-Meng, Hu Zhong-Yi. LDA feature selection based text classification and user clustering in Chinese online health community. *Journal of the China Society for Scientific and Technical Information*, 2017, **36**(11): 1183–1191
(吴江, 侯绍新, 靳萌萌, 胡忠义. 基于 LDA 模型特征选择的在线医疗社区文本分类及用户聚类研究. 情报学报, 2017, **36**(11): 1183–1191)
- 89 Sim J, Wright C C. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 2005, **85**(3): 257–268
- 90 Graham B. Kaggle Diabetic Retinopathy Detection Competition Report, University of Warwick, UK, 2015.
- 91 Li Man-Hua, Sun Hao-Peng, You Qi-Dong. Establishment and evaluation of prediction models for the discovery of CYP1A2 inhibitors. *Journal of China Pharmaceutical University*, 2013, **44**(5): 401–409
(李曼华, 孙昊鹏, 尤启冬. CYP1A2 抑制剂预测模型的建立及评价. 中国药科大学学报, 2013, **44**(5): 401–409)
- 92 Yang Bei. Bayesian Network Modeling: A Cohort Study of Cognitive Function Evaluation among the Elderly [Master Thesis], Shanxi Medical University, China, 2017.
(杨蓓. 贝叶斯网络模型在老年人认知功能评价队列研究中的应用 [硕士学位论文], 山西医科大学, 中国, 2017.)
- 93 Tian Xiu-Juan, Liu Guo-Cai, Gu Shan-Shan, Ju Zhong-Jian, Liu Jin-Guang, Gu Dong-Dong. Deep learning in medical image analysis and its challenges. *Acta Automatica Sinica*, 2018, **44**(3): 401–424
(田娟秀, 刘国才, 谷珊珊, 鞠忠建, 刘劲光, 顾冬冬. 医学图像分析深度学习方法研究与挑战. 自动化学报, 2018, **44**(3): 401–424)
- 94 Zheng T, Xie W, Xu L L, He X Y, Zhang Y, You M R, et al. A machine learning-based framework to identify type 2 diabetes through electronic health records. *International Journal of Medical Informatics*, 2017, **97**: 120–127
- 95 Zheng Guang-Yuan, Liu Xia-Bi, Han Guang-Hui. Survey on medical image computer aided detection and diagnosis systems. *Journal of Software*, 2018, **29**(5): 1471–1514
(郑光远, 刘峡壁, 韩光辉. 医学影像计算机辅助检测与诊断系统综述. 软件学报, 2018, **29**(5): 1471–1514)
- 96 Rahu M, McKee M. Epidemiological research labelled as a violation of privacy: The case of Estonia. *International Journal of Epidemiology*, 2008, **37**(3): 678–682
- 97 Tramèr F, Zhang F, Juels A, Reiter M K, Ristenpart T. Stealing machine learning models via prediction APIs. In: Proceedings of the 25th USENIX Conference on Security Symposium. Berkeley, CA, United States: USENIX Association, 2016. 601–618
- 98 Shokri R, Stronati M, Song C Z, Shmatikov V. Membership inference attacks against machine learning models. In: Proceedings of the 2017 IEEE Symposium on Security and Privacy (SP). San Jose, CA, USA: IEEE, 2017. 3–18
- 99 Abadi M, Chu A, Goodfellow I, McMahan H B, Mironov I, Talwar K, et al. Deep learning with differential privacy. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. Vienna, Austria: ACM, 2016. 308–318
- 100 Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608, 2017.
- 101 Goodman B, Flaxman S. European union regulations on algorithmic decision-making and a right to explanation. arXiv Preprint arXiv: 1606.08813, 2016.
- 102 Lum K, Isaac W. To predict and serve? *Significance*, 2016, **13**(5): 14–19
- 103 Faber J, Fonseca L M. How sample size influences research outcomes. *Dental Press Journal of Orthodontics*, 2014, **19**(4): 27–29



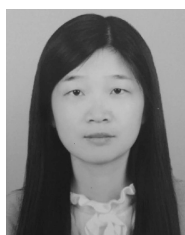
(and data mining.)

范家伟 北京邮电大学硕士研究生。主要研究方向为人工智能与数据挖掘。E-mail: jwfan@bupt.edu.cn
(**FAN Jia-Wei** Master student at Beijing University of Posts and Telecommunications. His research interest covers artificial intelligence



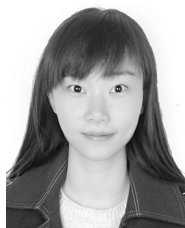
(processing.)

张如如 北京邮电大学博士研究生。主要研究方向为深度学习和医学影像处理。E-mail: zhangru@bupt.edu.cn
(**ZHANG Ru-Ru** Ph.D. candidate at Beijing University of Posts and Telecommunications. Her research interest covers deep learning and medical image processing.)



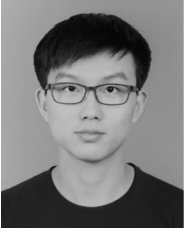
(version.)

陆萌 北京邮电大学硕士研究生。主要研究方向为数据挖掘和计算机视觉。E-mail: buptLumeng@bupt.edu.cn
(**LU Meng** Master student at Beijing University of Posts and Telecommunications. Her research interest covers data mining and computer version.)



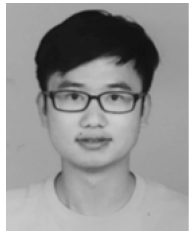
(image processing.)

何佳雯 北京邮电大学硕士研究生。主要研究方向为深度学习和图像处理。E-mail: euphy@bupt.edu.cn
(**HE Jia-Wen** Master student at Beijing University of Posts and Telecommunications. Her research interest covers deep learning and image processing.)



康霄阳 北京邮电大学硕士研究生. 主要研究方向为机器学习和计算机视觉.

E-mail: kangxiaoyang@bupt.edu.cn
(**KANG Xiao-Yang** Master student at Beijing University of Posts and Telecommunications. His research interest covers machine learning and computer vision.)



柴文俊 北京邮电大学硕士研究生. 主要研究方向为深度学习和计算机视觉.

E-mail: chaiwenjun@bupt.edu.cn
(**CHAI Wen-Jun** Master student at Beijing University of Posts and Telecommunications. His research interest covers deep learning and computer vision.)



石坤达 北京邮电大学硕士研究生. 主要研究方向为机器学习和计算机视觉. E-mail: cy.z.feng@gmail.com

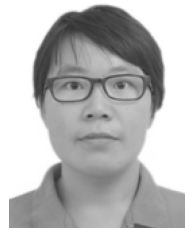
(**SHI Shen-Da** Master student at Beijing University of Posts and Telecommunications. His research interest covers machine learning and computer vision.)



宋美娜 教授, 教育部信息网络工程研究中心主任. 主要研究方向为服务计算, 云计算, 超大规模信息服务系统和人工智能. 本文通信作者.

E-mail: mnsong@gmail.com

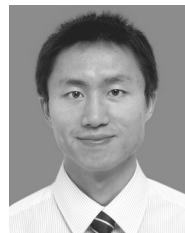
(**SONG Mei-Na** Ph.D., professor, Director of the Information Network Engineering Research Center of the Ministry of Education. Her research interest covers service computing, cloud computing, very large scale information service system, and artificial intelligence. Corresponding author of this paper.)



鄂海红 副教授, CCSA 移动互联网应用与终端技术委员会 WG1 副组长. 主要研究方向为移动互联网、大数据、云计算和人工智能.

E-mail: ehaihong@bupt.edu.cn

(**E Hai-Hong** Ph.D., associate professor, CCSA Mobile Internet Application and Terminal Technology Committee WG1 Deputy Leader. Her research interest covers mobile internet, big data, cloud computing, and artificial intelligence.)



欧中洪 副教授, 北京邮电大学计算机学院副院长. 主要研究方向为大数据分析、深度学习技术.

E-mail: zhonghong.ou@bupt.edu.cn

(**OU Zhong-Hong** Ph.D., associate professor, and Deputy Dean of the School of Computer Science, Beijing University of Posts and Telecommunications. His research interests cover big data analytics and deep learning techniques.)